

트위터에 나타난 국제 커뮤니케이션 네트워크의 하위집단 분석

Subgroup Analysis of Global Communication Network on Twitter

서일정, 조재희
광운대학교 경영대학 경영학부

Il-Jung Seo(iljung.seo@gmail.com), Jaehee Cho(mis1@kw.ac.kr)

요약

본 연구는 국제 커뮤니케이션 네트워크에 존재하는 국가들의 하위집단을 분석하여 국제 커뮤니케이션 현상에 대한 실증적 이해를 높이고자 하였다. 이를 위하여 전 세계 트위터 사용자의 커뮤니케이션 데이터를 수집하여 국제 커뮤니케이션 네트워크를 구성하였으며, 국제연합과 세계경제포럼에서 사용하는 국가의 지리적, 경제적 속성을 추가하였다. 국제 커뮤니케이션 네트워크에서 국가의 위치를 파악하기 위하여 중심성 분석과 핵심-주변 분석을 실시하였으며, 하위집단의 커뮤니케이션 관계를 이해하기 위하여 하위집단 내와 하위집단 간의 응집력을 분석하였다. 또한 국제 커뮤니케이션 네트워크에 내재되어 있는 하위집단을 발견하기 위하여 군집 분석을 실시하였다. 분석 결과에 의하면, 국제 커뮤니케이션 네트워크에는 핵심 국가들이 중심적인 위치를 차지하는 핵심-주변 구조가 있는 것으로 나타났으며 경제적 수준에 따른 국제 커뮤니케이션의 위계 구조가 존재하는 것으로 나타났다. 또한 군집 분석을 통해 경제, 지리, 역사, 문화, 종교 등 다양한 내용의 동질감을 갖는 하위집단을 발견하였다.

■ 중심어 : | 국제 커뮤니케이션 | 소셜 네트워크 | 네트워크 분석 | 트위터 |

Abstract

We investigated subgroups within a global communication network to improve the empirical understanding of global communication phenomenon from the social network perspective. We collected global communication data from Twitter and constructed a global communication network. We also added countries' geographic and economic properties used in the United Nations and the World Economic Forum. We analyzed the subgroups' structure within the global communication network using centrality analysis, core-peripheral analysis, and cohesion analysis. We also detected communities embedded within the global communication network with modularity-based community detection methods. We found that the core countries occupy central positions in the global communication network and there is a hierarchical communication structure among the economic subgroups. Furthermore, we discovered some communities within the global communication network and found that countries within the communities can have homophily such as economy, geography, history, culture, and religion.

■ keyword : | Global Communication | Social Network | Network Analysis | Twitter |

* 이 논문은 2014학년도 광운대학교 교내 학술연구비 지원에 의해 연구되었음

접수일자 : 2016년 02월 15일

심사완료일 : 2016년 03월 03일

수정일자 : 2016년 03월 02일

교신저자 : 조재희, e-mail : mis1@kw.ac.kr

I. 서론

커뮤니케이션의 관계론적 관점에 의하면, 인간의 커뮤니케이션은 상호 간의 이해에 도달하기 위해 정보를 생산하고 공유하는 역동적 과정이다[1]. 사람들은 정보 교환, 경제거래, 정치교류 등을 목적으로 커뮤니케이션을 하면서 사회적 관계를 맺는다[2]. 커뮤니케이션을 통해 맺어진 사회적 관계는 커뮤니케이션 네트워크를 구성하고[3], 사회적 구조를 형성한다[4]. 따라서 커뮤니케이션 네트워크를 분석함으로써 사회적 구조의 성격과 내용을 탐색할 수 있고 행위자의 상대적 위치와 역할을 파악할 수 있다[1].

커뮤니케이션 기술의 혁신은 국가의 지리적 경계를 넘어 국가 사이의 커뮤니케이션을 증가시켰다[5]. 이러한 국제 커뮤니케이션의 증가는 범세계적인 사회적 관계를 형성하면서 세계를 하나의 네트워크 사회로 만들고 있다[6]. 최근에 등장한 소셜 네트워크 서비스는 사용자 사이의 커뮤니케이션 뿐만 아니라 사회적 관계를 관리할 수 있는 기능을 제공함으로써 네트워크 사회로의 이행을 가속화시키고 있다.

국가 간 정보의 불균형 흐름[7], 정보중속과 문화중속[8], 글로벌라이제이션(globalization)[9] 등 국제 커뮤니케이션에 대한 다양한 이론적 논의가 있었다. 하지만, 국제 커뮤니케이션 연구의 가장 큰 문제점은 양적인 연구방법을 이용한 실증적 증거가 부족하다는 것이다[10]. 그 이유는 지금까지 국제 커뮤니케이션 현상을 분석할 수 있는 범세계적인 데이터 수집이 쉽지 않았기 때문이다. 하지만 대표적 소셜 네트워크 서비스인 트위터가 전 세계 트위터 사용자의 커뮤니케이션을 수집할 수 있는 수단을 제공하면서 국제 커뮤니케이션 네트워크뿐만 아니라 다양한 형태의 소셜 네트워크 분석 연구들이 가능해지고 있다[11][12].

이상의 논의를 바탕으로, 본 연구는 트위터에서 발생한 커뮤니케이션 데이터를 수집하여 국제 커뮤니케이션 네트워크의 구조를 탐색적으로 분석하고자 하였다. 일반적으로 네트워크 내 행위자의 동질감(homophily)이 서로 높으면 사회적 관계나 커뮤니케이션 관계가 형성될 가능성이 높아진다[13]. 이렇게 구성된 네트워크 내의 부분집단을 하위집단이라고 한다. 하위집단을 분

석하면 사회 구조에 관한 이론적 함의를 도출할 수 있다[14]. 따라서 본 연구에서는 국제 커뮤니케이션 네트워크 내에 존재하는 국가들의 하위집단을 분석하여 국제 커뮤니케이션 현상에 대한 실증적 이해를 높이고자 하였다.

II. 연구 방법

1. 데이터

국제 커뮤니케이션 네트워크를 구성하기 위하여 사용자의 위치 정보를 포함하고 있는 트윗(geotagged tweets)을 사용하였다. 트위터 사용자는 그들의 팔로워들에게 140자 미만의 메시지를 게시할 수 있고(트윗), 다른 사용자가 게시한 트윗을 팔로워들과 공유할 수 있다(리트윗). 또한 다른 사용자의 트윗에 직접적으로 회신할 수 있다(답글). 이러한 활동 중에서 ‘답글’은 다른 사람들과 개인적으로 커뮤니케이션을 하는 유일한 방식이기 때문에 본 연구에서는 트위터 사용자의 ‘답글’에 초점을 맞추었다.

트위터의 Public Streaming API[15]를 이용하여 2014년 9월 22일부터 동년 10월 23일까지 283,766,115 트윗을 수집하였다. 61개에 달하는 필드 중에서 연구 목적에 부합하는 3개의 필드만을 분석에 이용하였다. 그 필드는 답글을 보낸 사람의 식별자인 ‘user id’, 답글을 받은 사람의 식별자인 ‘in reply to user id’, 그리고 답글을 작성한 위치의 국가 코드인 ‘country code’이다.

개인 수준의 커뮤니케이션을 국가 수준의 국제 커뮤니케이션 네트워크로 변환하기 위하여 일련의 전처리 과정을 수행하였다. 첫째, ‘in reply to user id’와 ‘country code’가 널(null) 값이 아닌 트윗을 추출하였다. 둘째, 자기루프(self-loop)를 삭제하였다. 자기루프는 사용자가 자기 자신에게 답글한 트윗이다. 셋째, 사용자의 거주국가를 결정하였다. 거주국가는 사용자가 트윗을 자주하는 위치의 국가로 정의하고 빈도수가 가장 많은 국가를 거주국가로 결정하였다. 마지막으로, 개인 수준의 트윗을 ‘country code’를 이용하여 국가 수준으로 집계하였다. 그리고 개인 수준의 경우와 같이 자기루프를 삭제하였다. 다시 말해서 국가 내 커뮤니케이

션을 데이터에서 삭제하였다. 이러한 과정을 거쳐 231개 국가 사이의 11,468개 링크로 구성된 국제 커뮤니케이션 네트워크를 구성하였다.

지리적 하위집단의 커뮤니케이션 구조를 분석하기 위하여 국제연합(United Nations)에서 사용하는 지리적 지역 구분[16]을 국가의 속성으로 사용하였다. 국제연합은 지리적 지역을 대륙과 하위 지역으로 구분하고 있으며, 본 연구에서는 대륙을 지리적 속성으로 사용하였다. 대륙은 '아프리카', '아메리카', '아시아', '유럽', '오세아니아'로 구분된다. 그리고 국가의 경제적 구분을 위하여, 세계경제포럼(World Economic Forum)에서 개발한 개발단계[17]를 국가의 경제적 속성으로 사용하였다. 세계경제포럼은 개발단계를 '1단계', '1-2단계', '2단계', '2-3단계', '3단계'로 구분하고 있다.

2. 분석 방법

국제 커뮤니케이션 네트워크에서 국가의 위치와 역할을 분석하기 위하여 중심성 분석을 실시하였다. 본 연구에서는 대표적 중심성 지수인 내향연결 중심성, 외향연결 중심성, 근접 중심성, 매개 중심성, 위세 중심성을 측정하였다[18].

노드 i 의 연결 중심성은 식(1)과 같이 근접행렬 x 에서 노드 i 와 연결된 노드 j 의 총 개수로 측정된다.

$$D_i = \sum_j x_{ij} \quad (1)$$

연결 중심성은 링크의 방향에 따라 내향연결 중심성과 외향연결 중심성으로 구분할 수 있다. 내향연결 중심성은 노드로 들어오는 링크의 개수이며 명성으로 해석되고, 외향연결 중심성은 노드에서 나가는 링크의 개수이며 사교성으로 해석된다.

근접 중심성은 식(2)와 같이 노드 간 최단경로길이의 합으로 측정된다. d_{ij} 는 노드 i 에서 노드 j 까지 가는 최단경로의 길이이고 n 은 네트워크에 존재하는 노드의 개수라고 할 때, 노드 i 의 근접 중심성 C_i 는 최소 가능 값인 $n-1$ 을 모든 j 에 대한 최단경로길이의 합으로 나눈 역수로 측정된다. 근접 중심성이 높으면 커뮤니케이션의 효율성 또는 정보의 접근성이 높다고 해석한다.

$$C_i = \frac{n-1}{\sum_j d_{ij}} \quad (2)$$

매개 중심성은 두 노드 간 최단경로에 해당 노드가 존재하는 빈도로 정의된다. 매개 중심성을 구하는 방법은 식(3)과 같다. g_{ijk} 는 노드 j 를 지나면서 노드 i 와 노드 k 를 연결하는 최단경로의 개수이고, g_{ik} 는 노드 i 와 노드 k 를 연결하는 최단경로의 총 개수이다. 매개 중심성은 정보 흐름의 통제성으로 해석할 수 있다.

$$B_j = \sum_{i < k} \frac{g_{ijk}}{g_{ik}} \quad (3)$$

마지막으로 위세 중심성은 연결 중심성을 확장한 개념으로 해당 노드의 연결뿐만 아니라 근접한 노드들의 연결까지 고려한다. 위세 중심성 E_i 는 식(4)와 같이 해당 노드와 근접한 노드들의 연결 중심성을 모두 합한 값으로 산출된다. 이때, e_j 는 근접 노드의 위세 중심성 값이고 λ 는 비율 상수로 아이겐 값이다. 다시 말해서 각 노드의 위세 중심성은 해당 노드와 연결된 노드들의 위세 중심성에 비례한다.

$$E_i = \lambda \sum_j x_{ij} e_j \quad (4)$$

위세 중심성은 인기도 또는 영향력으로 해석된다. 위세 중심성이 높은 노드는 연결 중심성이 높은 노드들에 연결되어 있다는 것을 의미하며, 따라서 연결 중심성이 낮은 노드도 위세 중심성은 높을 수 있다.

하위집단의 응집력을 분석하기 위하여 각각의 하위집단별로 밀도, 평균연결, 평균경로길이, 근접계수를 산출하였다. 그리고 하위집단 간의 커뮤니케이션 관계를 분석하기 위하여 하위집단 간의 밀도를 산출하였다.

밀도는 집단 내에 존재하는 실제 링크의 개수를 최대 가능 링크의 개수로 나눈 값이고, 평균연결은 각 노드에 연결된 링크 개수의 평균이다. 그리고 평균경로길이는 모든 노드 간 최단경로길이의 평균이고, 근접계수는 각 노드의 자아 네트워크 밀도를 평균한 값이다.

국제 커뮤니케이션 네트워크의 핵심-주변 구조를 파악하기 위하여 UCINET[19]을 이용하여 핵심-주변 분

석을 실시하였다. 핵심-주변 분석은 블록 모델링 기법으로 노드들 사이의 높은 연결 밀도를 갖는 노드들을 핵심 집단으로 분류하고 낮은 연결 밀도를 갖는 노드들을 주변 집단으로 분류한다[18]. 또한 핵심-주변 구조를 직관적으로 이해하기 위해 네트워크를 이중 원 레이아웃으로 시각화하였다.

마지막으로, 국제 커뮤니케이션 네트워크에 내재되어 있는 군집(community)을 발견하기 위하여 루베인(Louvain) 알고리즘[20]을 이용하였다. 이 방법은 모듈 최적화를 기반으로 하는 휴리스틱 방법으로, 네트워크의 군집을 발견하는 여러 알고리즘과 비교하여 정확성과 효율성이 탁월한 것으로 평가되고 있다[21].

III. 분석 결과

1. 중심성 분석과 핵심-주변 분석

중심성 지수의 상위 40개 국가를 중심성 지수별로 [표 3]에 정리하였다. 내향연결 중심성과 외향연결 중심성은 실제 측정값이고, 근접 중심성과 매개 중심성, 그리고 위세 중심성은 정규화한 값이다. 중심성 지수 사이의 상관관계를 분석한 결과에 의하면 매개 중심성을 제외한 모든 지수의 상관이 상당히 높은 것으로 나타났다[표 1]. 이러한 결과는 [표 3]에서도 확인할 수 있는데 뉴질랜드, 케냐, 가나 등의 국가는 다른 중심성 순위에는 없고 매개 중심성 순위에만 올라왔다.

표 1. 중심성 지수 간 상관관계

중심성 지수	1	2	3	4	5
1. 내향연결 중심성					
2. 외향연결 중심성	0.99				
3. 근접 중심성	0.76	0.78			
4. 매개 중심성	0.61	0.61	0.48		
5. 위세 중심성	0.98	0.97	0.76	0.47	

핵심-주변 분석의 결과인 [표 2]를 보면, 핵심 국가 사이의 커뮤니케이션 밀도(0.91)는 주변 국가 사이의 커뮤니케이션 밀도(0.03)보다 높다. 그리고 핵심 국가로부터 주변 국가로 향하는 커뮤니케이션 밀도(0.25)

보다 주변 국가로부터 핵심 국가로 향하는 커뮤니케이션 밀도(0.27)가 높다. 이러한 결과는 핵심 국가들이 국제적 정보 흐름의 중심 위치를 차지하고 있어 정보 중심이 나타난다는 주장[8]을 지지하는 것이다.

표 2. 핵심-주변 국가의 커뮤니케이션 밀도

	핵심 국가	주변 국가
핵심 국가	0.91	0.25
주변 국가	0.27	0.03

국제 커뮤니케이션 네트워크의 핵심-주변 구조를 직관적으로 이해하기 위하여 [그림 1]처럼 이중 원으로 시각화하였다. 위세 중심성이 높은 40개 국가를 안쪽 원에 위치시키고, 낮은 국가를 바깥 원에 위치시켰다. 전체적인 구조를 명확히 파악하기 위하여 전체 11,468개 링크 중에서 가중치(링크 수)가 높은 상위 20%(2,294개)의 링크만 [그림 1]에 표현하였다. 핵심 국가 사이의 커뮤니케이션은 활발하지만 주변 국가 사이의 커뮤니케이션은 그렇지 못한 것을 확인할 수 있다.

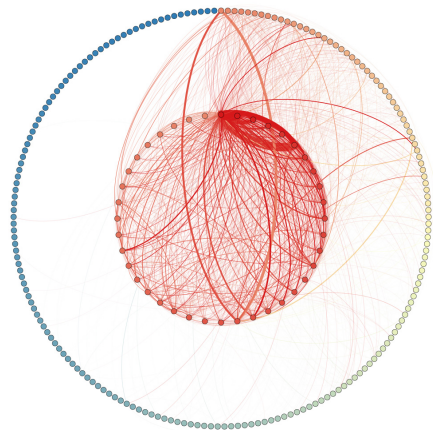


그림 1. 국가 간 커뮤니케이션 네트워크

2. 지리적 하위집단 분석

지리적 하위집단의 응집력 분석 결과에 의하면[표 4], 응집력이 가장 높은 집단은 유럽이고 응집력이 가장 낮은 집단은 오세아니아로 나타났다. 전체적으로 보면 유

표 3. 국가의 중심성 지수

내향연결 중심성	외향연결 중심성	근접 중심성	매개 중심성	위세 중심성					
미국	209	미국	205	미국	0.91	미국	0.14	미국	1.00
영국	195	영국	189	영국	0.86	영국	0.08	영국	0.99
캐나다	172	프랑스	169	프랑스	0.80	프랑스	0.04	캐나다	0.96
프랑스	167	캐나다	168	캐나다	0.80	캐나다	0.04	프랑스	0.95
이탈리아	157	브라질	152	브라질	0.75	브라질	0.03	이탈리아	0.93
브라질	157	스페인	149	스페인	0.75	스페인	0.03	독일	0.93
스페인	153	이탈리아	144	이탈리아	0.73	오스트레일리아	0.02	스페인	0.93
독일	151	독일	140	독일	0.72	이탈리아	0.02	브라질	0.93
터키	147	인도	140	인도	0.72	터키	0.02	터키	0.91
인도	142	오스트레일리아	139	오스트레일리아	0.72	독일	0.02	인도	0.90
인도네시아	140	네덜란드	138	네덜란드	0.72	포르투갈	0.02	네덜란드	0.89
오스트레일리아	139	터키	136	터키	0.72	남아공	0.02	필리핀	0.88
남아공	137	인도네시아	132	인도네시아	0.71	인도네시아	0.02	인도네시아	0.88
네덜란드	136	남아공	129	남아공	0.70	네덜란드	0.02	오스트레일리아	0.88
아르헨티나	135	필리핀	129	필리핀	0.70	아르헨티나	0.02	아르헨티나	0.87
필리핀	133	아르헨티나	128	아르헨티나	0.70	인도	0.01	일본	0.87
일본	132	러시아	127	러시아	0.70	일본	0.01	러시아	0.86
아일랜드	131	벨기에	121	벨기에	0.68	필리핀	0.01	벨기에	0.86
벨기에	130	일본	120	일본	0.68	멕시코	0.01	멕시코	0.86
멕시코	129	아일랜드	118	아일랜드	0.68	칠레	0.01	아일랜드	0.86
러시아	127	스위스	117	스위스	0.67	러시아	0.01	남아공	0.85
아랍에미리트	122	사우디아라비아	117	싱가포르	0.67	벨기에	0.01	아랍에미리트	0.84
스위스	120	나이지리아	116	나이지리아	0.67	나이지리아	0.01	스위스	0.83
말레이시아	120	베네수엘라	114	베네수엘라	0.67	아일랜드	0.01	말레이시아	0.82
사우디아라비아	115	아랍에미리트	113	아랍에미리트	0.67	스위스	0.01	싱가포르	0.81
스웨덴	113	스웨덴	112	스웨덴	0.66	모로코	0.01	스웨덴	0.81
베네수엘라	112	멕시코	110	멕시코	0.66	싱가포르	0.01	폴란드	0.79
콜롬비아	111	포르투갈	108	포르투갈	0.66	말레이시아	0.01	태국	0.79
이집트	109	이집트	107	이집트	0.66	이집트	0.01	베네수엘라	0.78
폴란드	108	말레이시아	106	말레이시아	0.65	베네수엘라	0.01	그리스	0.78
포르투갈	108	콜롬비아	105	콜롬비아	0.65	덴마크	0.01	이집트	0.78
태국	106	그리스	105	그리스	0.65	아랍에미리트	0.01	콜롬비아	0.78
그리스	105	파키스탄	104	파키스탄	0.65	뉴질랜드	0.01	노르웨이	0.77
싱가포르	105	태국	102	태국	0.65	콜롬비아	0.00	포르투갈	0.76
파키스탄	104	칠레	102	칠레	0.65	케냐	0.00	싱가포르	0.76
노르웨이	103	모로코	102	모로코	0.65	스웨덴	0.00	파키스탄	0.76
대한민국	102	폴란드	101	폴란드	0.64	파키스탄	0.00	칠레	0.75
나이지리아	101	덴마크	97	덴마크	0.64	폴란드	0.00	오스트리아	0.75
오스트리아	99	알제리	97	알제리	0.64	태국	0.00	대한민국	0.74
칠레	99	노르웨이	95	노르웨이	0.63	가나	0.00	덴마크	0.74

럽, 아시아, 아메리카가 아프리카, 오세아니아보다 상대적으로 응집력이 높은 것을 알 수 있다. 또한 좁은 세상

(small world)을 의미하는 평균경로길이는 유럽이 제일 짧고 아프리카가 제일 길다.

표 4. 지리적 하위집단의 속성

대륙	노드	링크	밀도	평균 연결	평균 경로길이	군집 계수
유럽	49	1,318	0.56	26.90	1.39	0.78
아시아	49	1,077	0.41	21.98	1.56	0.75
아메리카	54	827	0.28	15.32	1.71	0.67
아프리카	57	522	0.19	9.16	1.98	0.50
오세아니아	24	28	0.08	1.17	1.87	0.28
평균	47	754	0.03	14.91	1.70	0.60

[표 5]는 지리적 하위집단 간 커뮤니케이션 밀도를 행렬 형태로 표현한 것이다. 유럽과 아시아 사이의 커뮤니케이션 밀도(0.41)가 가장 높고, 오세아니아와 아프리카 사이의 커뮤니케이션 밀도(0.03)가 가장 낮다.

지역별로 살펴보면, 유럽은 아시아와 아메리카 사이의 밀도가 다른 지역과 비교하여 상대적으로 높고, 아시아는 유럽과 아메리카, 아메리카는 유럽과 아시아, 아프리카는 유럽과 아시아 사이의 밀도가 높다.

이상을 종합해 보면, 유럽은 국제 커뮤니케이션에서 중요한 위치를 차지하고 있다. 또한 유럽, 아시아, 아메리카를 잇는 연대가 존재하고 아프리카와 오세아니아는 이 연대로부터 분리되어 있다.

표 5. 지리적 하위집단 간 커뮤니케이션 밀도

	1	2	3	4	5
1. 유럽	0.56	0.41	0.27	0.17	0.08
2. 아시아	0.41	0.41	0.20	0.14	0.07
3. 아메리카	0.28	0.21	0.28	0.10	0.06
4. 아프리카	0.19	0.18	0.01	0.19	0.03
5. 오세아니아	0.08	0.08	0.07	0.03	0.08

3. 경제적 하위집단 분석

경제적 하위집단의 속성을 [표 6]에 정리하였다. 응집력을 노드 수에 영향을 덜 받는 군집계수와 평균연결로 평가하면, 응집력이 제일 높은 하위집단은 3단계의 집단이고 응집력이 제일 낮은 하위집단은 1-2단계의 집단이다. 밀도는 노드 수에 영향을 받는데, 1-2단계의 밀도가 1단계보다 높은 것은 1-2단계의 노드 수가 1단계보다 적기 때문이다.

표 6. 경제적 하위집단의 속성

개발단계	노드	링크	밀도	평균 연결	평균 경로길이	군집 계수
3단계	36	1,045	0.83	29.03	1.17	0.88
2-3단계	24	324	0.59	13.50	1.42	0.79
2단계	30	370	0.43	12.33	1.58	0.63
1-2단계	16	80	0.33	5.00	1.73	0.56
1단계	37	264	0.20	7.14	2.01	0.59
평균	29	417	0.48	13.40	1.58	0.69

[표 7]은 경제적 하위집단 간 커뮤니케이션 밀도를 행렬 형태로 표현한 것이다. 2-3단계에서 3단계로 향하는 커뮤니케이션 밀도가 제일 높고, 1-2단계와 1단계 사이의 커뮤니케이션 밀도가 제일 낮다.

해당 단계보다 상위 단계의 밀도가 하위 단계의 밀도보다 모든 단계에서 높다. 예를 들어, 2단계에서 밖으로 향하는 커뮤니케이션의 밀도는 3단계가 가장 높고 2-3단계, 1-2단계, 1단계 순으로 낮아진다.

이상을 종합해 보면, 경제적 개발정도와 집단의 응집력 사이에는 높은 관계가 있음을 알 수 있다. 다시 말해서 개발단계가 높을수록 다른 집단과의 커뮤니케이션 밀도가 높아진다.

표 7. 경제적 하위집단 간 커뮤니케이션 밀도

	1	2	3	4	5
1. 3단계	0.83	0.71	0.61	0.50	0.35
2. 2-3단계	0.73	0.59	0.53	0.44	0.26
3. 2단계	0.62	0.54	0.43	0.37	0.22
4. 1-2단계	0.54	0.44	0.41	0.33	0.19
5. 1단계	0.40	0.29	0.24	0.19	0.20

4. 군집 분석

국제 커뮤니케이션 네트워크에 존재하는 군집을 발견하기 위하여, 루베인 알고리즘[20]을 구현한 Gephi[22]를 사용하여 군집 분석을 실시하였다(randomize; weight; resolution=1).

군집 분석을 실시한 결과, [그림 2]와 같이 7개(modularity=0.25)의 군집을 발견하였다. 군집 A에는 미국, 브라질, 영국, 남아프리카 공화국, 오스트레일리

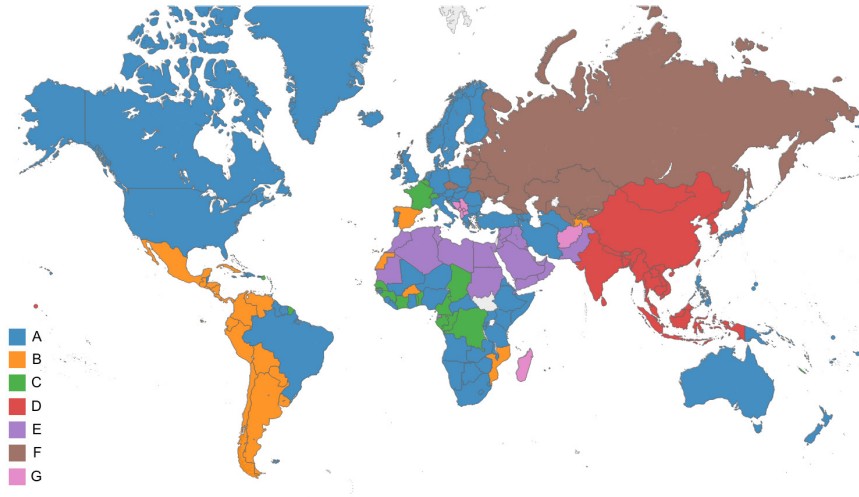


그림 2. 국제 커뮤니케이션 네트워크의 군집

아, 일본 등 다양한 대륙의 국가들이 포함되어 있다. 군집 B는 스페인과 스페인의 식민지였던 멕시코, 아르헨티나, 칠레 등 라틴 아메리카의 국가들을 포함하고 있다. 군집 C에는 프랑스와 프랑스의 식민지였던 카메룬, 토고, 기니 등의 국가들과 스위스, 벨기에 등의 국가들이 포함되어 있다. 군집 D는 중국, 인도, 인도네시아, 대한민국 등 아시아 국가들을 포함하고 있다. 군집 E는 아랍에미리트, 사우디아라비아, 이집트 등 서남 아시아와 북 아프리카에 위치한 이슬람 국가들이 포함되어 있다. 군집 F는 러시아와 구소련이었던 카자흐스탄, 우즈베키스탄 등의 국가들과 체코, 폴란드 등 동유럽 국가들을 포함하고 있다. 마지막으로 군집 G에는 세르비아, 코소보, 알바니아 등 발칸 반도에 위치한 국가들이 포함되어 있다.

표 8. 군집의 속성

개발단계	노드	링크	밀도	평균 연결	평균 경로길이	군집 계수
A	112	2,350	0.19	20.99	1.87	0.75
B	29	335	0.41	11.55	1.50	0.68
C	26	175	0.27	6.73	1.72	0.62
D	24	269	0.49	11.21	1.53	0.74
E	21	313	0.75	14.91	1.26	0.86
F	13	96	0.62	7.39	1.33	0.71
G	8	31	0.56	3.88	1.39	0.78

표 9. 군집 간 커뮤니케이션 밀도

	A	B	C	D	E	F	G
A	0.19	0.20	0.12	0.23	0.25	0.20	0.20
B	0.20	0.41	0.14	0.26	0.26	0.22	0.22
C	0.13	0.15	0.27	0.14	0.21	0.11	0.11
D	0.21	0.25	0.12	0.49	0.33	0.30	0.15
E	0.25	0.27	0.20	0.38	0.75	0.29	0.32
F	0.22	0.22	0.10	0.26	0.27	0.62	0.28
G	0.20	0.22	0.10	0.19	0.29	0.25	0.56

군집계수로 군집의 응집력을 비교해보면, 군집 E가 가장 높고 군집 C가 가장 낮다[표 8]. 군집 E는 [표 9]에서 보는 것과 같이 군집 내 응집력뿐만 아니라 다른 군집들과의 커뮤니케이션 밀도도 상대적으로 높다.

IV. 결론

본 연구는 국제 커뮤니케이션 현상의 실증적 이해를 높이고자 하였다. 연구 목적을 달성하기 위하여 전 세계 트위터 사용자의 커뮤니케이션 데이터를 수집하여 일련의 과정을 거쳐 국제 커뮤니케이션 네트워크를 구성하였다. 국제 커뮤니케이션 네트워크를 대상으로 중심성 분석, 핵심-주변 분석, 하위집단의 응집력 분석, 하위집단 간의 커뮤니케이션 밀도 분석, 군집 분석을 실

시하여 하위집단의 구조를 파악하였다.

연구의 결과를 요약하면 다음과 같다.

첫째, 핵심 국가들이 국제 커뮤니케이션 네트워크의 중심적인 위치를 차지하고 있다. 중심성 분석의 결과에 의하면 중심성 지수 사이의 상관성이 매우 높은 것으로 나타났다[표 1]. 이러한 결과는 여러 국가들과 활발히 커뮤니케이션을 하는 국가들이 명성, 사교성, 정보의 접근성, 정보 흐름의 통제성, 영향력 등 다양한 측면에서 국제 커뮤니케이션 네트워크의 중심적 위치를 차지하고 있다는 것을 의미한다. 또한 핵심-주변 분석을 통해 국제 커뮤니케이션 네트워크에 핵심-주변 구조가 존재하는 것을 확인하였다[표 2].

둘째, 국제 커뮤니케이션 관계 형성에 지리적 요인보다는 경제적 요인이 더 강하게 작용한다. 지리적 하위집단의 응집력과 경제적 하위집단의 응집력을 비교해보면[표 4][표 6], 경제적 하위집단의 응집력이 지리적 하위집단의 응집력보다 대체적으로 강한 것으로 나타났다. 예를 들어 지리적 하위집단에서 응집력이 제일 높은 유럽은 군집계수가 0.78이고, 경제적 하위집단에서 응집력이 제일 높은 3단계는 군집계수가 0.88이다.

셋째, 국가의 경제적 수준에 따라 국제 커뮤니케이션의 위계 구조가 존재한다[표 7]. 경제적 수준이 가장 높은 국가들(개발단계가 3단계인 국가들)은 동일 수준의 국가들 사이의 커뮤니케이션 관계가 제일 강하고 경제적 수준이 낮아질수록 관계의 정도가 약해진다. 그리고 경제적 수준이 낮은 국가들은 본국과 비교하여 경제적 수준이 낮은 국가들보다 경제적 수준이 높은 국가들과의 커뮤니케이션 관계가 강하다.

마지막으로, 국제 커뮤니케이션 네트워크에는 경제, 지리, 역사, 종교 등 다양한 내용의 동질감을 갖는 하위집단이 존재한다. 군집 분석의 결과에 의하면[그림 2], 가장 많은 국가를 포함하면서 경제적 연대로 추정되는 군집 A, 과거 제국주의의 역사적 배경을 공유하는 군집 B와 C, 지리적 또는 문화적 동질감을 갖는 군집 D, 종교적 동질감을 갖는 군집 E, 구소련과 관련된 군집 F 등을 확인할 수 있다.

본 연구는 민간 부문의 국제 커뮤니케이션 네트워크에 내재되어 있는 다양한 내용의 동질감을 갖는 국가

하위집단을 밝혔다. 동질적 하위집단 내 국가들은 규범과 목적을 공유하면서 특정 사안에 대한 유사한 태도와 행동을 보이며, 이질적 하위집단에 대한 부정적 행태를 가질 수 있다. 따라서 본 연구에서 밝힌 하위집단의 결과를 국제교류, 국제경영 등에 관심이 높은 단체 또는 기관에서 활용한다면 국제관계 전략을 수립하는데 도움이 될 것이다.

서론에서 밝힌 바와 같이 지금까지 국제 커뮤니케이션 연구는 현상에 대한 다양한 이론적 논의가 활발하였지만, 경험적 데이터를 분석한 실증적 연구가 부족하였다. 본 연구는 범세계적인 커뮤니케이션 데이터를 수집하여 국제 커뮤니케이션 네트워크를 분석함으로써, 국제 커뮤니케이션에 대한 실증적 연구의 가능성을 확인하였다.

특정 소셜 네트워크 서비스를 이용한 사용자의 커뮤니케이션을 대상으로 하는 단일 연구의 분석 결과를 국제 커뮤니케이션의 일반적인 현상이라고 주장하기에는 무리가 있을 수 있다. 이를 극복하기 위하여 국가별 트위터 이용 현황, 인터넷 보급률 등과 같은 다양한 요소들을 추가하여 분석한다면, 연구 결과의 견고성을 확보할 수 있을 것이다. 그리고 전 세계 국가들에 대한 역사 및 문화에 대한 넓고 깊은 이해를 바탕으로, 군집 분석의 과정과 해석을 정교하게 수행한다면 국제 커뮤니케이션에 대한 좀 더 풍부하고 의미 있는 통찰을 얻을 수 있을 것이다.

참 고 문 헌

- [1] E. M. Rogers and D. L. Kincaid, *Communication Networks: Toward a New Paradigm for Research*, New York: The Free Press, 1981.
- [2] 김경모, “커뮤니케이션 연결망 분석의 이론적 기초에 관한 탐색적 접근,” *커뮤니케이션 이론*, 제1권, 제2호, pp.162-207, 2005.
- [3] I. de Sola Pool, “Communication systems,” In I. de Sola Pool et al., *Handbook of Communication*, Chicago: Rand McNally College Publishing

- Company, pp.3-26, 1973.
- [4] K. Krippendorff, "Communication and the Genesis of Structure," *General Systems*, Vol.16, pp.171-185, 1971.
- [5] 김승현, "세계화와 국제 커뮤니케이션연구의 새로운 문제들," *한국언론정보학보*, 통권14호, pp.89-114, 2000.
- [6] M. Castells, *The Internet Galaxy: Reflections on the Internet, Business, and Society*, Oxford University Press, 2002.
- [7] J. Galtung and R. C. Vincent, *Global Glasnost: Toward a New World Information and Communication Order*, Hampton Press, 1992.
- [8] H. I. Schiller, *Information Inequality*, Routledge, 1995.
- [9] R. Robertson, "Glocalization: Time-Space and Homogeneity-Heterogeneity," In M. Featherstone et al., *Global Modernities*, Sage, pp.25-44, 1995.
- [10] 황상재, "주요 저널을 통해서 본 국제커뮤니케이션 연구동향과 시사점," *한국언론학회*, 봄철 정기학술 대회, pp.111-120, 2000.
- [11] 유병국, 김순홍, "소셜네트워크 분석을 통한 마케팅 전략," *한국콘텐츠학회논문지*, 제13권, 제5호, pp.396-407, 2013.
- [12] 조성은, 박한우, "정부부청의 소셜미디어 소통방식: 국가간 트위터 이용 및 연결망에 대한 탐색적 연구," *한국콘텐츠학회논문지*, 제13권, 제8호, pp.160-170, 2013.
- [13] E. M. Rogers and D. K. Bhowmik, "Homophily-Heterophily: Relational Concepts for Communication Research," *Public Opinion Quarterly*, Vol.34, No.4, pp.523-538, 1970.
- [14] P. Boracich, "Communication Networks and Collective Action," *Social Networks*, Vol.9, No.4, pp.389-396, 1987.
- [15] <http://dev.twitter.com/streaming/public>
- [16] <http://millenniumindicators.un.org/unsd/methods/m49/m49regin.htm>
- [17] <http://reports.weforum.org/global-competitiveness-report-2014-2015/>
- [18] S. P. Borgatti, M. G. Everett, and J. C. Johnson,

Analyzing Social Networks, London: SAGE Publications Limited, 2013.

- [19] <http://sites.google.com/site/ucinetsoftware/>
- [20] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast Unfolding of Communities in Large Networks," *Journal of Statistical Mechanics: Theory and Experiment*, Vol.10, P10008, 2008.
- [21] A. Lancichinetti and S. Fortunato, "Community Detection Algorithms: A Comparative Analysis," *Physical Review E*, Vol.80, No.5, 056117, 2009.
- [22] <http://gephi.org>

저 자 소 개

서 일 정(II-Jung Seo)

정회원



- 1999년 2월 : 경기대학교 경영정보학과(경영학사)
- 2006년 2월 : 광운대학교 경영정보학과(경영학석사)
- 2010년 2월 : 광운대학교 경영정보학과(경영학박사)

- 2011년 ~ 2012년 : 경북대학교 경상대학 초빙교수
 - 2016년 ~ 현재 : 광운대학교 경영대학 외래강사
- <관심분야> : 네트워크 분석, 데이터 시각화

조 재 희(Jaehee Cho)

정회원



- 1983년 2월 : 연세대학교 경영대학(경영학사)
- 1985년 8월 : 미국 마이애미대학교 경영대학원(경영학석사)
- 1991년 12월 : 미국 네브라스카대학교 경영대학(경영학박사)

- 1994년 ~ 현재 : 광운대학교 경영대학 교수
 - 2001년 : 미국 버클리대학교 교환교수
 - 2011년 ~ 2013년 : 한국BI데이터마이닝학회장
- <관심분야> : Twitter data analytics, 다차원분석, 네트워크분석, 데이터시각화